Model reduction for a class of singularly perturbed stochastic differential equations : Fast variable approximation (Extended Version)

Narmada Herath* and Domitilla Del Vecchio †

This is an extended version of a paper of the same title accepted to American Control Conference (ACC) 2016 [1].

Abstract

We consider a class of stochastic differential equations in singular perturbation form, where the drift terms are linear and diffusion terms are nonlinear functions of the state variables. In our previous work, we approximated the slow variable dynamics of the original system by a reduced-order model when the singular perturbation parameter ϵ is small. In this work, we obtain an approximation for the fast variable dynamics. We prove that the first and second moments of the approximation are within an $O(\epsilon)$-neighborhood of the first and second moments of the fast variable of the original system. The result holds for a finite time-interval after an initial transient has elapsed. We illustrate our results with a biomolecular system modeled by the chemical Langevin equation.

1 Introduction

Systems with multiple time-scales can be written in singular perturbation form, where the dynamics are separated into slow and fast, with a small parameter ϵ capturing the separation in time-scales. The analysis of singularly perturbed systems consists of obtaining a reduced-order model that approximates the dynamics of the system when the time-scale separation is large. In the deterministic setting, the main method used to obtain the reduced system is given by Tikhonov's theorem, which gives a set of reduced-order differential equations that approximate the slow variable dynamics, and a set of algebraic equations that approximate the fast variable dynamics [2, 3]. Another method that is used to analyze systems with multiple time-scales is the averaging principle, which gives a reduced-order model that approximates the dynamics of the slow variables [4].

*Narmada Herath is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 77 Mass. Ave, Cambridge MA nherath@mit.edu

†Domitilla Del Vecchio is with the Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Mass. Ave, Cambridge MA ddv@mit.edu

Singular perturbation techniques have also been developed for stochastic systems with multiple time-scales [5, 6, 3, 7]. However, these methods are not applicable to systems where the diffusion term of the fast variable is nonlinear, and is of the order $\sqrt{\epsilon}$, as seen, for example, in biomolecular systems modeled by the chemical Langevin equation [8]. Averaging methods for stochastic differential equations have also been developed, and they can be applied to the case where the diffusion term is on the order of $\sqrt{\epsilon}$ [9, 4]. However, the averaging methods only provide an approximation to the slow variable dynamics.

In our previous work, we considered a class of singularly perturbed stochastic differential equations with linear drift terms and nonlinear diffusion terms [10]. We obtained a reduced-order model that approximates the slow variable dynamics of the original system and we proved that the first and second moments of the reduced system are within an $O(\epsilon)$-neighborhood of the first and second moments of the original system, for a finite time interval. This result was extended in [11] to prove that all the moments of the reduced system are within an $O(\epsilon)$-neighborhood of the corresponding moments of the original system, for a finite time interval. Although, the above works provide an approximation to the slow variable dynamics, in many applications it is necessary to obtain an approximation for the fast variable in order to analyze the statistical properties of the system, such as the mean and the variance. In fact, in many biomolecular applications, the variables in the system may be affected by both slow and fast reactions, and thus the system is represented in singular perturbation form after using a coordinate transformation [12]. Therefore, to analyze the properties of the variables of interest, it is necessary to have a fast variable approximation. For example, in deterministic models of gene transcriptional networks, singular perturbation approach with both slow and fast variable approximations has been used to characterize the change in dynamics of transcription factors, caused by the interconnections with other components [13, 14]. Hence, in this work, we obtain an approximation for the fast variable dynamics of the original system in the form of an algebraic expression of the reduced slow variable dynamics. We prove that the first and second moments of the approximation are within an $O(\epsilon)$-neighborhood of the first and second moment dynamics of the fast variable of the original system. The result holds for a finite time interval, after a short transient has elapsed.

This paper is organized as follows. In Section II, we introduce the class of systems considered. In Section III, we define the fast variable approximation and derive the moments of the approximation. In Section IV, we prove that the first and second moments of the fast variable approximation are within an $O(\epsilon)$-neighborhood of those of the original system. The results are illustrated with an example in Section V.

2 System Model

Consider the singularly perturbed stochastc system

$$\dot{x} = f_x(x, z, t) + \sigma_x(x, z, t)\Gamma_x, \qquad x(0) = x_0 \tag{1}$$

$$\epsilon\dot{z} = f_z(x, z, t, \epsilon) + \sigma_z(x, z, t, \epsilon)\Gamma_z, \; z(0) = z_0 \tag{2}$$

where $x \in D_x \subset \mathbb{R}^n$ and $z \in D_z \subset \mathbb{R}^m$ are the slow and fast variables, respectively. Γ_x is a d_x-dimensional white noise process and let Γ_f be a d_f-dimensional white noise process. Then, Γ_z is a $(d_x + d_f)$-dimensional white noise process. We assume that the system (1) - (2) satisfies the following assumptions.

Assumption 1. The functions $f_x(x, z, t)$ and $f_z(x, z, t, \epsilon)$ are affine functions of the state variables x and z, i.e., we can write $f_x(x, z, t) = A_1x + A_2z + A_3(t)$, where $A_1 \in \mathbb{R}^{n\times n}$, $A_2 \in \mathbb{R}^{n\times m}$ and $A_3(t) \in \mathbb{R}^n$ and $f_z(x, z, t, \epsilon) = B_1x + B_2z + B_3(t) + \alpha(\epsilon)(B_4x + B_5z + B_6(t))$, where $B_1, B_4 \in \mathbb{R}^{m\times n}$, $B_2, B_5 \in \mathbb{R}^{m\times m}$, $B_3(t), B_6(t) \in \mathbb{R}^m$. Also, we have that $A_3(t)$ and $B_3(t)$ are continuously differentiable functions, and $\alpha(\epsilon)$ is a continuously differentiable function with $\alpha(0) = 0$.

Assumption 2. The matrix-valued functions $\sigma_x(x, z, t)$ and $\sigma_z(x, z, t, \epsilon)$ are such that, there exist continuously differentiable functions $\Phi(x, z, t) : \mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^n \to \mathbb{R}^{n\times n}$, $\Lambda(x, z, t, \epsilon) : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{m\times m}$, $\Theta(x, z, t, \epsilon) : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{m\times n}$, that satisfy

$$\sigma_x(x, z, t)\sigma_x(x, z, t)^T = \Phi(x, z, t), \tag{3}$$

$$\sigma_z(x, z, t, \epsilon)\sigma_z(x, z, t, \epsilon)^T = \epsilon\Lambda(x, z, t, \epsilon), \tag{4}$$

$$\sigma_z(x, z, t, \epsilon)\begin{bmatrix} \sigma_x(x, z, t) & 0 \end{bmatrix}^T = \Theta(x, z, t, \epsilon), \tag{5}$$

where the elements of $\Phi(x, z, t)$, $\Lambda(x, z, t, \epsilon)$, $\Theta(x, z, t, \epsilon)$ are affine functions of x and z, i.e., we can write $\mathbb{E}[\Phi(x, z, t)] = \Phi(\mathbb{E}[x], \mathbb{E}[z], t)$, $\mathbb{E}[\Lambda(x, z, t, \epsilon)] = \Lambda(\mathbb{E}[x], \mathbb{E}[z], t, \epsilon)$, and $\mathbb{E}[\Theta(x, z, t, \epsilon)] = \Theta(\mathbb{E}[x], \mathbb{E}[z], t, \epsilon)$. Also, we have that $\lim_{\epsilon\to 0} \Lambda(x, z, t, \epsilon) < \infty$ and $\lim_{\epsilon\to 0} \Theta(x, z, t, \epsilon) = 0$ for all x, z and t.

Assumption 3. *The matrix B_2 is Hurwitz.*

We assume that a unique, well-defined solution exists for the system (1) - (2), for a finite time-interval. The sufficient conditions for existence of a unique solution is given in [15], which consists of Lipschitz continuity and bounded growth conditions. The class of systems that satisfy Assumption 2, includes the case where the diffusion term is a square-root function of the state variables that may not satisfy the Lipschitz continuity condition. In this case, the sufficient conditions in [16] can be used to verify the existence of a unique solution.

Original System

$$\dot{x} = f_x(x,z,t) + \sigma_x(x,z,t)\ \Gamma_x$$
$$\epsilon\dot{z} = f_z(x,z,t,\epsilon) + \sigma_z(x,z,t,\epsilon)\ \Gamma_z$$

$\longrightarrow$

Reduced System

$$\dot{\bar{x}} = f_x(\bar{x},\gamma_1(\bar{x},t),t) + \sigma_x(\bar{x},\gamma_1(\bar{x},t),t)\ \Gamma_x$$

$\downarrow$ $\qquad$ $\xrightarrow{\epsilon \to 0}$ $\qquad$ $\downarrow$

Moments of the Original System

$$\frac{d}{dt}\begin{pmatrix} \mathbb{E}[x] \\ \mathbb{E}[xx^T] \\ \epsilon\mathbb{E}[z] \\ \epsilon\mathbb{E}[zz^T] \\ \epsilon\mathbb{E}[zx^T] \end{pmatrix} = \begin{pmatrix} g_1(\mathbb{E}[x],\mathbb{E}[z],t) \\ g_2(\mathbb{E}[x],\mathbb{E}[xx^T],\mathbb{E}[z],\mathbb{E}[zx^T],t) \\ g_3(\mathbb{E}[x],\mathbb{E}[z],t,\epsilon) \\ g_4(\mathbb{E}[x],\mathbb{E}[xx^T],\mathbb{E}[z],\mathbb{E}[zz^T],\mathbb{E}[zx^T],t,\epsilon) \\ g_5(\mathbb{E}[x],\mathbb{E}[xx^T],\mathbb{E}[z],\mathbb{E}[zz^T],\mathbb{E}[zx^T],t,\epsilon) \end{pmatrix}$$

Moments of the Reduced System

$$\begin{pmatrix} \frac{d}{dt}\mathbb{E}[\bar{x}] \\ \frac{d}{dt}\mathbb{E}[\bar{x}\bar{x}^T] \end{pmatrix} = \begin{pmatrix} g_1(\mathbb{E}[\bar{x}],\gamma_1(\mathbb{E}[\bar{x}],t),t) \\ g_2(\mathbb{E}[\bar{x}],\mathbb{E}[\bar{x}\bar{x}^T],\gamma_1(\mathbb{E}[\bar{x}],t),\gamma_2(\mathbb{E}[\bar{x}],\mathbb{E}[\bar{x}\bar{x}^T],t),t) \end{pmatrix}$$

Figure 1: Setting $\epsilon = 0$ in the moment dynamics of the original system yields the moment dynamics of the reduced system.

3 Fast variable approximation

In [10], we defined the reduced system which was shown to approximate the slow variable dynamics, as

$$\dot{\bar{x}} = f_x(\bar{x},\gamma_1(\bar{x},t),t) + \sigma_x(\bar{x},\gamma_1(\bar{x},t),t)\Gamma_x,\ \ \bar{x}(0) = x_0, \tag{6}$$

where

$$\gamma_1(x,t) = -B_2^{-1}(B_1x + B_3(t)), \tag{7}$$

is the solution to $f_z(x,z,t,0) = B_1x + B_2z + B_3(t) = 0$. We then quantified the error in the slow variable approximation using the first and the second moments and obtained the results summarized here in Theorem 1. First, consider the following function definitions for suitable constant matrices $a \in \mathbb{R}^n$, $b \in \mathbb{R}^{n\times n}$, $c \in \mathbb{R}^m$, $d \in \mathbb{R}^{m\times m}$, and $e \in \mathbb{R}^{m\times n}$.

$$\begin{aligned}
g_1(a,\gamma_1(a,t),t) &= A_1a + A_2\gamma_1(a,t) + A_3(t), \\
g_2(a,b,\gamma_1(a,t),\gamma_2(a,b,t),t) &= A_1b + A_2\gamma_2(a,b,t) + bA_1^T + A_3(t)a^T + \gamma_2(a,b,t)^TA_2^T + aA_3(t)^T \\
&\quad + \Phi(a,\gamma_1(a,t),t), \\
g_3(a,c,t,\epsilon) &= B_1a + B_2c + B_3(t) + \alpha(\epsilon)(B_4a + B_5c + B_6(t)), \\
g_4(a,b,c,d,e,t,\epsilon) &= eB_1^T + dB_2^T + cB_3(t)^T + \alpha(\epsilon)(eB_4^T + dB_5^T + cB_6(t)^T) + B_1e^T \\
&\quad + \Lambda(a,c,t,\epsilon) + B_2d + B_3(t)c^T + \alpha(\epsilon)(B_4e^T + B_5d + B_6(t)c^T), \\
g_5(a,b,c,d,e,t,\epsilon) &= \epsilon(eA_1^T + dA_2^T + cA_3(t)^T) + B_1b + B_2e + B_3(t)a^T + \Theta(a,c,t,\epsilon) \\
&\quad + \alpha(\epsilon)(B_4e^T + B_5d + B_6(t)c^T),
\end{aligned}$$

Theorem 1. *Consider the original system in (1) - (2) and the reduced system in (6). Under Assumptions 1 - 3, the commutative diagram in Fig. 1 holds. Furthermore, there exists $t_1 > 0$ such that*

$$\begin{aligned}
&\|\mathbb{E}[\bar{x}(t)] - \mathbb{E}[x(t)]\| = O(\epsilon),\ \ t \in [0,t_1], \\
&\|\mathbb{E}[\bar{x}(t)\bar{x}(t)^T] - \mathbb{E}[x(t)x(t)^T]\|_F = O(\epsilon),\ \ t \in [0,t_1],
\end{aligned}$$

where $\|.\|_F$ *is the Frobenius norm.*

In [10] we illustrate via an example that, although $\gamma_1(\bar{x}(t), t)$ provides a good approximation when it is used in the slow variable dynamics, it does not provide a good approximation of the fast variable, in contrast to the deterministic case. To illustrate this further, consider the following example as in [10].

$$\dot{x} = -a_1 x + a_2 z, \tag{8}$$
$$\dot{z} = -z + v_1\sqrt{\epsilon}\Gamma, \tag{9}$$

where $a_1 > 0$. We have that $z = \gamma_1(x, t) = 0$, and the reduced system is given by

$$\dot{\bar{x}} = -a_1\bar{x}.$$

Calculating the steady state second moment dynamics of x and z in the system (8) - (9), we obtain

$$\mathbb{E}[x^2] = \frac{a_2^2 v_2^2 \epsilon}{2a_1(1 + a_1\epsilon)}, \tag{10}$$
$$\mathbb{E}[z^2] = \frac{v_1^2}{2}. \tag{11}$$

We also have that the $\mathbb{E}[\bar{x}^2] = 0$ at steady state and $\mathbb{E}[\gamma_1(\bar{x}(t), t)^2] = 0$. Therefore, we see that when $\epsilon = 0$, $\|\mathbb{E}[x^2] - \mathbb{E}[\bar{x}^2]\| = 0$ but $\|\mathbb{E}[z^2] - \mathbb{E}[\gamma_1(\bar{x}, t)^2]\| = \frac{v_1^2}{2}$. Thus, $z(t) = \gamma(\bar{x}(t), t)$ does not approximate the fast variable dynamics well.

In this work, we seek an approximation to the fast variable in the form

$$\bar{z}(t) = \gamma_1(\bar{x}(t), t) + g(\bar{x}(t), t)N, \tag{12}$$

where $N \in \mathbb{R}^d$ is a random vector, whose components are independent standard normal random variables that are also independent of $\bar{x}$, and $g(\bar{x}(t), t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{m\times d}$ is a suitable function. We call equation (12), *the reduced fast system*. We aim at determining the function $g(\bar{x}(t), t)$ such that the first and second moments of $z(t)$ in (1) - (2) are well approximated by the first and second moments of $\bar{z}(t)$ in (12). To this end, define the functions ψ, γ_2 and γ_3 for $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^{n\times n}$ such that

$$\psi(a, t) = \int_0^\infty e^{(B_2 v)}\Lambda(a, \gamma_1(a, t), t, 0)e^{(B_2^T v)}dv, \tag{13}$$
$$\gamma_2(a, b, t) = -B_2^{-1}(B_1 b + B_3(t)a^T), \tag{14}$$
$$\gamma_3(a, b, t) = -\gamma_2(a, b, t)B_1^T(B_2^{-1})^T - \gamma_1(a, t)B_3(t)^T(B_2^{-1})^T + \psi(a, t). \tag{15}$$

We now make the following claim:

Claim 1. *Let $g(x(t),t)$ satisfy the Lyapunov equation*

$$g(\bar{x}(t),t)g(\bar{x}(t),t)^T B_2^T + B_2 g(\bar{x}(t),t)g(\bar{x}(t),t)^T = -\Lambda(\bar{x},\gamma_1(\bar{x}(t),t),t,0). \tag{16}$$

Then, the first and second moments of $\bar{z}(t)$ defined in (12) can be written in the form

$$\mathbb{E}[\bar{z}(t)] = \gamma_1(\mathbb{E}[\bar{x}(t)],t), \tag{17}$$

$$\mathbb{E}[\bar{z}(t)\bar{z}(t)^T] = \gamma_3(\mathbb{E}[\bar{x}(t)],\mathbb{E}[\bar{x}(t)\bar{x}(t)^T],t). \tag{18}$$

Proof. From equation (12), we have

$$\mathbb{E}[\bar{z}] = \mathbb{E}[\gamma_1(\bar{x},t) + g(\bar{x},t)N].$$

Employing the linearity of the expectation operator and of the function $\gamma_1(\bar{x},t)$, we have

$$\mathbb{E}[\bar{z}] = \gamma_1(\mathbb{E}[\bar{x}],t) + \mathbb{E}[g(\bar{x},t)N].$$

Since N is a random vector whose components are standard normal random variables, we have that $\mathbb{E}[N]=0$. As the function $g(\bar{x},t)$ and the random vector N are independent, we further obtain

$$\mathbb{E}[\bar{z}] = \gamma_1(\mathbb{E}[\bar{x}],t).$$

Similarly, the second moment of the reduced fast variable can be calculated using (12) as

$$\begin{aligned}
\mathbb{E}[\bar{z}\bar{z}^T] &= \mathbb{E}[(\gamma_1(\bar{x},t) + g(\bar{x},t)N)(\gamma_1(\bar{x},t) + g(\bar{x},t)N)^T], \\
&= \mathbb{E}[\gamma_1(\bar{x},t)\gamma_1(\bar{x},t)^T] + \mathbb{E}[\gamma_1(\bar{x},t)N^T g(\bar{x},t)^T] \\
&\quad + \mathbb{E}[g(\bar{x},t)N\gamma_1(\bar{x},t)^T] + \mathbb{E}[g(\bar{x},t)NN^T g(\bar{x},t)^T].
\end{aligned} \tag{19}$$

Let $G(\bar{x},t) = g(\bar{x},t)N\gamma_1(\bar{x},t)^T$ and write $\gamma_1(\bar{x},t) = [\gamma_1(\bar{x},t)_1,\ldots,\gamma_1(\bar{x},t)_m]^T$, $g(\bar{x},t) = [g(\bar{x},t)_{ij}]$ for $i \in \{1,m\}$, $j \in \{1,d\}$ and $N = [N_1,\ldots,N_d]^T$. Then, the entries of $G(\bar{x},t)$ can be written as

$$G(\bar{x},t)_{ik} = \sum_{j=1}^{d} \gamma_1(\bar{x},t)_k g(\bar{x},t)_{ij} N_j, \qquad \text{where} \qquad i \in \{1,m\} \text{ and } k \in \{1,m\}.$$

As the functions $\gamma_1(\bar{x},t)$ and $g(\bar{x},t)$ are independent from N, taking the expectation yields

$$\mathbb{E}[G(\bar{x},t)_{ik}] = \sum_{j=1}^{d} \mathbb{E}[\gamma_1(\bar{x},t)_k g(\bar{x},t)_{ij}]\mathbb{E}[N_j].$$

Since N_j is a standard normal random variable we have that $\mathbb{E}[N_j] = 0$ for all j. Therefore, we have $\mathbb{E}[G(\bar{x}, t)] = 0$. Similarly, we have that $\gamma_1(\bar{x}, t)N^T g(\bar{x}, t)^T = G(\bar{x}, t)^T$, and $\mathbb{E}[G(\bar{x}, t)^T] = 0$. Thus, we obtain

$$\mathbb{E}[\gamma_1(\bar{x}, t)N^T g(\bar{x}, t)^T] = \mathbb{E}[g(\bar{x}, t)N\gamma_1(\bar{x}, t)^T]^T = 0. \tag{20}$$

Let $H(\bar{x}, t) = g(\bar{x}, t)NN^T g(\bar{x}, t)^T$. The entries of $H(\bar{x}, t)$ can be expressed as

$$H(\bar{x}, t)_{ik} = \sum_{l=1}^{d} \sum_{j=1}^{d} g(\bar{x}, t)_{il} g(\bar{x}, t)_{kj} N_j N_l.$$

Since $\gamma_1(\bar{x}, t)$ and $g(\bar{x}, t)$ are independent from N, taking the expectation, we have

$$\mathbb{E}[H(\bar{x}, t)_{ik}] = \sum_{l=1}^{d} \sum_{j=1}^{d} \mathbb{E}[g(\bar{x}, t)_{il} g(\bar{x}, t)_{kj}]\mathbb{E}[N_j N_l].$$

As N_j are independent standard normal random variables, we have that $\mathbb{E}[N_l N_j] = \mathbb{E}[N_l]\mathbb{E}[N_j] = 0$, for all $l \neq j$. When $l = j$, we have that $N_l N_j = N_j^2$, which yields $\mathbb{E}[N_j^2] = var(N_j) = 1$. Then, we obtain

$$\mathbb{E}[H(\bar{x}, t)_{ik}] = \sum_{j=1}^{d} \mathbb{E}[g(\bar{x}, t)_{ij} g(\bar{x}, t)_{kj}],$$

which results in

$$\mathbb{E}[H(\bar{x}, t)] = \mathbb{E}[g(\bar{x}, t)NN^T g(\bar{x}, t)^T] = \mathbb{E}[g(\bar{x}, t)g(\bar{x}, t)^T]. \tag{21}$$

Substituting (20) and (21) into (19) yields

$$\mathbb{E}[\bar{z}\bar{z}^T] = \mathbb{E}[\gamma_1(\bar{x}, t)\gamma_1(\bar{x}, t)^T] + \mathbb{E}[g(\bar{x}, t)g(\bar{x}, t)^T]. \tag{22}$$

From (16), we have that

$$\mathbb{E}[g(\bar{x}, t)g(\bar{x}, t)^T]B_2^T + B_2\mathbb{E}[g(\bar{x}, t)g(\bar{x}, t)^T] = -\Lambda(\mathbb{E}[\bar{x}], \gamma_1(\mathbb{E}[\bar{x}], t), t, 0).$$

for which, under Assumption 3, the unique solution is given by

$$\mathbb{E}[g(\bar{x}, t)g(\bar{x}, t)^T] = \int_0^\infty e^{(B_2 v)}\Lambda(\mathbb{E}[\bar{x}], \gamma_1(\mathbb{E}[\bar{x}], t), t, 0)e^{(B_2^T v)} dv = \psi(\mathbb{E}[\bar{x}], t), \tag{23}$$

where ψ is defined in (13).

Therefore, using that $\gamma_1(x, t) = -B_2^{-1}(B_1 x + B_3(t))$, and substituting (23) into (21) leads to

$$\mathbb{E}[\bar{z}\bar{z}^T] = \mathbb{E}[(-B_2^{-1}(B_1\bar{x} + B_3(t)))(-B_2^{-1}(B_1\bar{x} + B_3(t)))^T] + \psi(\mathbb{E}[\bar{x}], t),$$

which, using (14) and (15) can be rewritten as

$$\begin{aligned}\mathbb{E}[\bar{z}\bar{z}^T] &= -\gamma_2(\mathbb{E}[\bar{x}], \mathbb{E}[\bar{x}\bar{x}^T], t)B_1^T(B_2^{-1})^T - \gamma_1(\mathbb{E}[\bar{x}], t)B_3(t)^T(B_2^{-1})^T + \psi(\mathbb{E}[\bar{x}], t),\\ &= \gamma_3(\mathbb{E}[\bar{x}], \mathbb{E}[\bar{x}\bar{x}^T], t).\end{aligned}$$

□

From Theorem 1 we have that the first and second moment dynamics for the fast variables of the original system $\mathbb{E}[z(t)]$, $\mathbb{E}[z(t)z(t)^T]$ and the dynamics of the mixed moments $\mathbb{E}[z(t)x(t)^T]$ are given by

$$\epsilon\frac{d\mathbb{E}[z]}{dt} = g_3(\mathbb{E}[x], \mathbb{E}[z], t, \epsilon), \tag{24}$$

$$\epsilon\frac{d\mathbb{E}[zz^T]}{dt} = g_4(\mathbb{E}[x], \mathbb{E}[xx^T], \mathbb{E}[z], \mathbb{E}[zz^T], \mathbb{E}[zx^T], t, \epsilon), \tag{25}$$

$$\epsilon\frac{d\mathbb{E}[zx^T]}{dt} = g_5(\mathbb{E}[x], \mathbb{E}[xx^T], \mathbb{E}[z], \mathbb{E}[zz^T], \mathbb{E}[zx^T], t, \epsilon). \tag{26}$$

Now, we analyze the moment dynamics for the fast variable of the original system, that is $\mathbb{E}[z(t)]$ and $\mathbb{E}[z(t)z(t)^T]$, when $\epsilon = 0$.

Claim 2. *Setting $\epsilon = 0$, in the moment dynamics (24) and (26), results in*

$$\mathbb{E}[z(t)] = \gamma_1(\mathbb{E}[x(t)], t), \tag{27}$$

$$\mathbb{E}[z(t)z(t)^T] = \gamma_3(\mathbb{E}[x(t)], \mathbb{E}[x(t)x(t)^T], t). \tag{28}$$

Proof. Setting $\epsilon = 0$ in (24) - (25), we obtain

$$B_1\mathbb{E}[x] + B_2\mathbb{E}[z] + B_3(t) = 0, \tag{29}$$

$$B_1\mathbb{E}[xx^T] + B_2\mathbb{E}[zx^T] + B_3(t)\mathbb{E}[x^T] = 0. \tag{30}$$

Under Assumption 3, we have that the unique solutions to equations (29) - (30) is given by

$$\mathbb{E}[z] = -B_2^{-1}(B_1\mathbb{E}[x] + B_3(t)) = \gamma_1(\mathbb{E}[x], t), \tag{31}$$

$$\begin{aligned}\mathbb{E}[zx^T] &= -B_2^{-1}(B_1\mathbb{E}[xx^T] + B_3(t)\mathbb{E}[x^T])\\ &= \gamma_2(\mathbb{E}[x], \mathbb{E}[xx^T], t).\end{aligned} \tag{32}$$

We have that equation (31) is in the form of equation (27), proving the first equality.

Setting $\epsilon = 0$ in (25), we obtain

$$\mathbb{E}[zz^T]B_2^T + B_2\mathbb{E}[zz^T] = -\mathbb{E}[zx^T]B_1^T - \mathbb{E}[z]B_3(t)^T - B_1\mathbb{E}[xz^T] - B_3(t)\mathbb{E}[z^T] - \Lambda(\mathbb{E}[x], \mathbb{E}[z], t, 0).$$

Using the expressions $\mathbb{E}[z] = \gamma_1(\mathbb{E}[x], t)$ and $\mathbb{E}[zx^T] = \gamma_2(\mathbb{E}[x], \mathbb{E}[xx^T], t)$ from (31) - (32), we have that

$$\begin{aligned}\mathbb{E}[zz^T]B_2^T + B_2\mathbb{E}[zz^T] = &-\gamma_2(\mathbb{E}[x], \mathbb{E}[xx^T], t)B_1^T - \gamma_1(\mathbb{E}[x], t)B_3(t)^T - B_1\gamma_2(\mathbb{E}[x], \mathbb{E}[xx^T], t)^T \\ &- B_3(t)\gamma_1(\mathbb{E}[x], t)^T - \Lambda(\mathbb{E}[x], \gamma_1(\mathbb{E}[x], t), t, 0). \end{aligned} \tag{33}$$

The equation (33) takes the form of the Lyapunov equation,

$$A^T P + PA = -Q,$$

with

$$\begin{aligned} &P = \mathbb{E}[zz^T], \\ &Q(\mathbb{E}[x], \mathbb{E}[xx^T], t) = \gamma_2(\mathbb{E}[x], \mathbb{E}[xx^T], t)B_1^T + \gamma_1(\mathbb{E}[x], t)B_3(t)^T + B_1\gamma_2(\mathbb{E}[x], \mathbb{E}[xx^T], t)^T \\ &\qquad\qquad + B_3(t)\gamma_1(\mathbb{E}[x], t)^T + \Lambda(\mathbb{E}[x], \gamma_1(\mathbb{E}[x], t), t, 0), \\ &A = B_2^T. \end{aligned}$$

Therefore, under Assumption 3, there exists a unique solution $\mathbb{E}[zz^T] = h(\mathbb{E}[x], \mathbb{E}[xx^T], t)$, to equation (33). To prove that $h(\mathbb{E}[x], \mathbb{E}[xx^T], t) = \gamma_3(\mathbb{E}[x], \mathbb{E}[xx^T], t)$, we substitute (15) into (33) (noting that $\gamma_3(\mathbb{E}[x], \mathbb{E}[xx^T], t)$ is symmetric), and simplify further to obtain

$$\begin{aligned} &-\gamma_2(\mathbb{E}[x], \mathbb{E}[xx^T], t)B_1^T - \gamma_1(\mathbb{E}[x], t)B_3(t)^T + \psi(\mathbb{E}[x], t)B_2^T - B_1\gamma_2(\mathbb{E}[x], \mathbb{E}[xx^T], t)^T \\ &- B_3(t)\gamma_1(\mathbb{E}[x], t)^T + B_2\psi(\mathbb{E}[x], t) \\ &= -\gamma_2(\mathbb{E}[x], \mathbb{E}[xx^T])B_1^T - \gamma_1(\mathbb{E}[x], t)B_3(t)^T - B_1\gamma_2(\mathbb{E}[x], \mathbb{E}[xx^T], t)^T - B_3(t)\gamma_1(\mathbb{E}[x], t)^T \\ &- \Lambda(\mathbb{E}[x], \gamma_1(\mathbb{E}[x], t), t, 0). \end{aligned}$$

Canceling the common terms on both sides finally yields

$$\psi(\mathbb{E}[x], t)B_2^T + B_2\psi(\mathbb{E}[x], t) = -\Lambda(\mathbb{E}[x], \gamma_1(\mathbb{E}[x], t), t, 0). \tag{34}$$

From the definition of ψ in (13), we have that $\psi(\mathbb{E}[x], t) = \int_0^\infty e^{(B_2 v)}\Lambda(\mathbb{E}[x], \gamma_1(\mathbb{E}[x], t), t, 0)e^{(B_2^T v)}dv$, which is the unique solution to the Lyapunov equation in (34), under Assumption 3. Therefore we have shown that $\mathbb{E}[zz^T] = \gamma_3(\mathbb{E}[x], \mathbb{E}[xx^T], t)$ is the unique solution to (33). $\square$

4 Main Result

In this section, we quantify the error between the moments of the fast variable of the original system given by (24) - (25) and moments of the reduced fast system given by (17) - (18). To this end, we have the following Lemma, which is an extension to the results in the commutative diagram in Fig. 1.

$$
\begin{array}{ccc}
\begin{array}{c}
\text{Original System}\\
\dot{x} = f_x(x,z,t) + \sigma_x(x,z,t)\ \Gamma_x\\
\epsilon\dot{z} = f_z(x,z,t,\epsilon) + \sigma_z(x,z,t,\epsilon)\ \Gamma_z
\end{array}
& \longrightarrow &
\begin{array}{c}
\text{Reduced System}\\
\dot{\bar{x}} = f_x(\bar{x},\gamma_1(\bar{x},t),t) + \sigma_x(\bar{x},\gamma_1(\bar{x},t),t)\ \Gamma_x\\
\boxed{\bar{z} = \gamma_1(\bar{x},t) + g(\bar{x},t)N}
\end{array}\\
\downarrow & & \downarrow\\
\text{Moments of the Original System} & \xrightarrow{\epsilon \to 0} & \text{Moments of the Reduced System}
\end{array}
$$

$$
\frac{d}{dt}\begin{pmatrix} \mathbb{E}[x]\\ \mathbb{E}[xx^T]\\ \epsilon\mathbb{E}[z]\\ \epsilon\mathbb{E}[zz^T]\\ \epsilon\mathbb{E}[zx^T] \end{pmatrix} = \begin{pmatrix} g_1(\mathbb{E}[x],\mathbb{E}[z],t)\\ g_2(\mathbb{E}[x],\mathbb{E}[xx^T],\mathbb{E}[z],\mathbb{E}[zx^T],t)\\ g_3(\mathbb{E}[x],\mathbb{E}[z],t,\epsilon)\\ g_4(\mathbb{E}[x],\mathbb{E}[xx^T],\mathbb{E}[z],\mathbb{E}[zz^T],\mathbb{E}[zx^T],t,\epsilon)\\ g_5(\mathbb{E}[x],\mathbb{E}[xx^T],\mathbb{E}[z],\mathbb{E}[zz^T],\mathbb{E}[zx^T],t,\epsilon) \end{pmatrix}
\qquad
\begin{array}{c}
\begin{pmatrix} \frac{d}{dt}\mathbb{E}[\bar{x}]\\ \frac{d}{dt}\mathbb{E}[\bar{x}\bar{x}^T] \end{pmatrix} = \begin{pmatrix} g_1(\mathbb{E}[\bar{x}],\gamma_1(\mathbb{E}[\bar{x}],t),t)\\ g_2(\mathbb{E}[\bar{x}],\mathbb{E}[\bar{x}\bar{x}^T],\gamma_1(\mathbb{E}[\bar{x}],t),\gamma_2(\mathbb{E}[\bar{x}],\mathbb{E}[\bar{x}\bar{x}^T],t),t) \end{pmatrix}\\
\boxed{\begin{pmatrix} \mathbb{E}[\bar{z}]\\ \mathbb{E}[\bar{z}\bar{z}^T] \end{pmatrix} = \begin{pmatrix} \gamma_1(\mathbb{E}[\bar{x}],t)\\ \gamma_3(\mathbb{E}[\bar{x}],\mathbb{E}[\bar{x}\bar{x}^T],t) \end{pmatrix}}
\end{array}
$$

Figure 2: Setting $\epsilon = 0$ in the moment dynamics for the fast variable of the original system yields the moment dynamics of the reduced fast system.

Lemma 1. *Under Assumptions 1 - 3, the commutative diagram in Fig. 2 holds.*

Proof. Proof follows from Theorem 1, Claim 1 and Claim 2. □

Theorem 2. Consider the original system in (1) - (2) and the reduced fast system in (12). Under Assumptions 1 - 3 there exists $\epsilon^* > 0,\ t_1,\ t_b > 0$ with $t_1 > t_b$ such that for $\epsilon < \epsilon^*$ we have

$$
\begin{aligned}
&\|\mathbb{E}[\bar{z}(t)] - \mathbb{E}[z(t)]\| = O(\epsilon),\ t \in [t_b, t_1],\\
&\|\mathbb{E}[\bar{z}(t)\bar{z}(t)^T] - \mathbb{E}[z(t)z(t)^T]\|_F = O(\epsilon),\ t \in [t_b, t_1],
\end{aligned} \tag{35}
$$

where $\|.\|_F$ is the Frobenius norm.

Proof. As the moment dynamics are deterministic, the results in (35) can be proven by applying the Tikhonov's theorem to the moment dynamics of the original system and moment dynamics of the reduced system given in Fig. 2. To this end, we first prove that the assumptions of the Tikhonov's theorem in [2] are satisfied.

In order to ensure the global exponential stability of the boundary layer dynamics for the moments of the original system, we define the error variables

$$
\begin{aligned}
b_1 &:= \mathbb{E}[z] - \gamma_1(\mathbb{E}[x], t),\\
b_2 &:= \mathbb{E}[zx^T] - \gamma_2(\mathbb{E}[x], \mathbb{E}[xx^T], t),\\
b_3 &:= \mathbb{E}[zz^T] - \gamma_3(\mathbb{E}[x], \mathbb{E}[xx^T], t).
\end{aligned} \tag{36}
$$

Letting $\tau := t/\epsilon$ denote the time variable in the fast time scale, it was shown in [10] that the boundary layer dynamics of b_1 and b_2, given by

$$\frac{db_1}{d\tau} = B_2 b_1,$$
$$\frac{db_2}{d\tau} = B_2 b_2,$$

are globally exponentially stable under Assumption 3. Next we analyze the boundary layer dynamics of b_3. From (36), the derivative of the variable b_3 with respect to time t is given by

$$\epsilon\frac{db_3}{dt} = \epsilon\frac{d\mathbb{E}[zz^T]}{dt} - \epsilon\frac{\partial\gamma_3(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial t} - \epsilon\frac{\partial\gamma_3(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial\mathbb{E}[x]}\frac{d\mathbb{E}[x]}{dt}$$
$$- \epsilon\frac{\partial\gamma_3(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial\mathbb{E}[xx^T]}\frac{d\mathbb{E}[xx^T]}{dt}. \tag{37}$$

Writing $\tau = t/\epsilon$, and using equation (25), the dynamics of b_3 in the fast timescale τ are given by

$$\frac{db_3}{d\tau} = g_4(\mathbb{E}[x],\mathbb{E}[xx^T],\mathbb{E}[z],\mathbb{E}[zz^T],\mathbb{E}[zx^T],t,\epsilon) - \frac{\partial\gamma_3(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial\mathbb{E}[x]}\frac{d\mathbb{E}[x]}{d\tau} - \frac{\partial\gamma_3(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial\tau}$$
$$- \frac{\partial\gamma_3(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial\mathbb{E}[xx^T]}\frac{d\mathbb{E}[xx^T]}{d\tau}, \tag{38}$$

in which we have from Fig. 2, that the moments of the original system satisfy

$$\frac{d\mathbb{E}[x]}{d\tau} = \epsilon g_1(\mathbb{E}[x],\mathbb{E}[z],t), \tag{39}$$
$$\frac{d\mathbb{E}[xx^T]}{d\tau} = \epsilon g_2(\mathbb{E}[x],\mathbb{E}[xx^T],\mathbb{E}[z],\mathbb{E}[zx^T],t). \tag{40}$$

Using equation (15), we have that

$$\frac{\partial\gamma_3(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial\tau} = \epsilon B_2^{-1}\frac{dB_3(t)}{dt}\mathbb{E}[x^T]B_1^T(B_2^{-1})^T + \epsilon B_2^{-1}\frac{dB_3(t)}{dt}B_3(t)^T(B_2^{-1})^T$$
$$- \epsilon\gamma_1(\mathbb{E}[x],t)\frac{dB_3(t)}{dt}(B_2^{-1})^T + \epsilon\frac{\partial}{\partial t}\int_0^\infty e^{(B_2 v)}\Lambda(\mathbb{E}[x],\gamma_1(\mathbb{E}[x],t),t,0)e^{(B_2^T v)}dv. \tag{41}$$

The boundary layer dynamic for b_3 are given by setting $\epsilon = 0$ in (38) and using $\mathbb{E}[z] = b_1 + \gamma_1(\mathbb{E}[x],t)$, $\mathbb{E}[zx^T] = b_2 + \gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)$, and $\mathbb{E}[zz^T] = b_3 + \gamma_3(\mathbb{E}[x],\mathbb{E}[xx^T],t)$. Due to the linearity of the functions g_1, g_2, g_3, g_4, g_5 in the commutative diagram of Fig. 2, we have that the solutions $\mathbb{E}[x], \mathbb{E}[z], \mathbb{E}[xx^T], \mathbb{E}[zx^T]$, $\mathbb{E}[zz^T]$ exist and are bounded on a finite time interval $t \in [0, t_1]$ for some finite $t_1 > 0$. Therefore, setting $\epsilon = 0$ in (39) and (40) we have that $\frac{d\mathbb{E}[x]}{d\tau} = 0$ and $\frac{d\mathbb{E}[xx^T]}{d\tau} = 0$.

Under Assumption 1, we have that $\frac{dB_3(t)}{dt}$ is continuous in t and therefore it is bounded on the finite time interval $t \in [0, t_1]$. Furthermore, we observe that, the matrix multiplications in the term $e^{(B_2 v)}\Lambda(\mathbb{E}[x],\gamma_1(\mathbb{E}[x],t),t,0)e^{(B_2^T v)}$ appearing in equation (41) results in linear combinations of the entries of Λ. Therefore, due to the continuous differentiability of Λ with respect to its arguments under Assumption 1, we have that the expression $\frac{\partial}{\partial t}\int_0^\infty e^{(B_2 v)}\Lambda(\mathbb{E}[x],\gamma_1(\mathbb{E}[x],t),t,0)e^{(B_2^T v)}dv$ is continuous with t. Hence, it is bounded on the finite time interval $t\in[0,t_1]$. Equation (41) thus implies that $\frac{\partial\gamma_3(\mathbb{E}[x],\mathbb{E}[xx^T],t)}{\partial\tau}=0$ when $\epsilon=0$. Therefore, setting $\epsilon=0$ in (38), using that $\alpha(0)=0$ and $\Lambda(x,z,t,0)<\infty$ from Assumption 1 and Assumption 2, and taking $\mathbb{E}[z]=b_1+\gamma_1(\mathbb{E}[x],t)$, $\mathbb{E}[zx^T]=b_2+\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t)$, $\mathbb{E}[zz^T]=b_3+\gamma_3(\mathbb{E}[x],\mathbb{E}[xx^T],t)$, we obtain the dynamics for b_3 as

$$\begin{aligned}\frac{db_3}{d\tau}&=(b_2+\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t))B_1^T+(b_3+\gamma_3(\mathbb{E}[x],\mathbb{E}[xx^T],t))B_2^T+(b_1+\gamma_1(\mathbb{E}[x],t))B_3(t)^T\\&\quad+B_1(b_2+\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t))^T+\Lambda(\mathbb{E}[x],\mathbb{E}[z],t,0)\\&\quad+B_2(b_3+\gamma_3(\mathbb{E}[x],\mathbb{E}[xx^T],t))^T+B_3(t)(b_1+\gamma_1(\mathbb{E}[x],t))^T.\end{aligned}$$

Substituting here the expression of γ_3 from (15), yields

$$\begin{aligned}\frac{db_3}{d\tau}&=(b_2+\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t))B_1^T+(b_3-\gamma_2(\mathbb{E}[\bar{x}],\mathbb{E}[\bar{x}\bar{x}^T],t)B_1^T(B_2^{-1})^T-\gamma_1(\mathbb{E}[\bar{x}],t)B_3(t)^T(B_2^{-1})^T\\&\quad+\psi(\mathbb{E}[x],t))B_2^T+(b_1+\gamma_1(\mathbb{E}[x],t))B_3(t)^T+B_1(b_2+\gamma_2(\mathbb{E}[x],\mathbb{E}[xx^T],t))^T\\&\quad+B_2(b_3-\gamma_2(\mathbb{E}[\bar{x}],\mathbb{E}[\bar{x}\bar{x}^T],t)B_1^T(B_2^{-1})^T-\gamma_1(\mathbb{E}[\bar{x}],t)B_3(t)^T(B_2^{-1})^T+\psi(\mathbb{E}[x],t))^T\\&\quad+B_3(t)(b_1+\gamma_1(\mathbb{E}[x],t))^T+\Lambda(\mathbb{E}[x],\gamma_1(\mathbb{E}[x],t),t,0).\end{aligned}\tag{42}$$

From (13), we have that $\psi(\mathbb{E}[x],t)=\int_0^\infty e^{(B_2 v)}\Lambda(\mathbb{E}[x],\gamma_1(\mathbb{E}[x],t),t,0)e^{(B_2^T v)}dv$, which is the unique solution to the Lyapunov equation

$$\psi(\mathbb{E}[x],t)B_2^T+B_2\psi(\mathbb{E}[x],t)=-\Lambda(\mathbb{E}[x],\gamma_1(\mathbb{E}[x],t),t,0).\tag{43}$$

Therefore using (43) in (42), we finally obtain

$$\frac{db_3}{d\tau}=b_3B_2^T+B_2b_3^T+b_2B_1^T+B_1b_2^T+B_3(t)b_1^T+b_1B_3(t)^T.\tag{44}$$

Under Assumption 3, we have that the matrix B_2 is Hurwitz and therefore the dynamics of b_1 and b_2 are globally exponentially stable. Then, using the solution of (44) for b_3 given by [17]

$$\begin{aligned}b_3(\tau)=e^{B_2\tau}b_3(0)e^{B_2^T\tau}+\int_0^\tau e^{B_2(\tau-v)}&(b_2(v)B_1^T+B_1b_2(v)^T+B_3(t)b_1(v)^T\\&+b_1(v)B_3(t)^T)(e^{B_2(\tau-v)})^T dv,\end{aligned}$$

it follows that there exists a positive constants C_1 and r_1 such that $\|b_3(\tau)\|_F\le C_1(\|b_1(0)\|_F+\|b_2(0)\|_F+\|b_3(0)\|_F)e^{-r_1\tau}$, where $\|.\|_F$ denotes the Frobenius norm. Then, taking $Y=[b_1\ b_2\ b_3]$, and considering the exponential stability of b_1 and b_2, we can write $\|Y(\tau)\|_F\le C\|Y(0)\|_F e^{-r\tau}$ for positive constants C and r. Therefore, we have that the origin is a globally exponentially stable equilibrium point of the boundary layer dynamics.

Under Assumption 1 we have that the functions g_1, g_2, g_3, g_4, g_5 and their first partial derivatives with respect to their arguments are continuous. We also have that the functions $\gamma_1(\mathbb{E}[x], t)$, $\gamma_2(\mathbb{E}[x], \mathbb{E}[xx^T], t)$, $\gamma_3(\mathbb{E}[x], \mathbb{E}[xx^T], t)$ have continuous first partial derivatives with respect to their arguments. Due to the linearity of the functions g_1 and g_2, the reduced system has a unique solution for $t \in [0, t_1]$. Therefore, the assumptions of the Tikhonov's theorem are satisfied. Applying the Tikhonov's theorem to the moment dynamics of the original system, we then obtain the result in (35). $\square$

4.1 Illustrative Example

Next, we consider again the motivating example in Section 3, and approximate the fast variable dynamics of the original system by the reduced fast system $\bar{z}(t) = \gamma_1(\bar{x}(t), t) + g(\bar{x}(t), t)N$ given in (12).

Setting $\epsilon = 0$, we obtain $\gamma_1(\bar{x}(t), t) = 0$. To obtain $g(\bar{x}(t), t)$, we solve the equation

$$g(\bar{x}(t), t)g(\bar{x}(t), t)^T(-1) + (-1)g(\bar{x}(t), t)g(\bar{x}(t), t)^T = -v_1^2$$

which yields $g(\bar{x}(t), t) = \sqrt{v_1/2}$. Therefore, the reduced fast system is given by

$$\bar{z}(t) = \frac{v_1}{\sqrt{2}}N.$$

We have that $\mathbb{E}[\bar{z}(t)^2] = \frac{v_1^2}{2}\mathbb{E}[N^2]$, where $\mathbb{E}[N^2] = 1$. From equation (11), we also have that the steady state second moment of the fast variable dynamics of the original system (9) is given by $\mathbb{E}[z^2] = \frac{v_1^2}{2}$. Therefore, it follows that $\|\mathbb{E}[z^2] - \mathbb{E}[\bar{z}^2]\| = 0$ and that the reduced fast system provides a good approximation for the fast variable dynamics of the original system.

5 Example

In this section, we apply the results to a biomolecular system that exhibits time-scale separation. Consider the system given in Fig. 3, where the transcription factor X binds to the promoter p_2 and regulates the production of protein G, while also binding to a non-regulatory binding site p_1. It has been show that the amount of non-regulatory binding sites - also referred to as decoy sites - can alter the speed and the shape of the response of protein X [18, 19, 20]. Stochastic effects of this system have also been studied in [21], using the chemical Master equation. In this section, we model the dynamics of the system using the chemical Langevin equation and obtain a reduced model, taking into account the time-scale separation in the system.

The chemical reactions for the system can be written as follows: $\phi \underset{\delta_1}{\overset{k(t)}{\rightleftharpoons}} X$, $X + p_1 \underset{k_{\text{off1}}}{\overset{k_{\text{on1}}}{\rightleftharpoons}} C_1$, $X + p_2 \underset{k_{\text{off2}}}{\overset{k_{\text{on2}}}{\rightleftharpoons}} C_2$, $\mathrm{C}_2 \xrightarrow{\beta} \mathrm{C}_2 + \mathrm{G}$, $\mathrm{G} \xrightarrow{\delta} \phi$, where $k(t)$ is the production rate of X, $k_{\text{on1}}, k_{\text{off1}}$

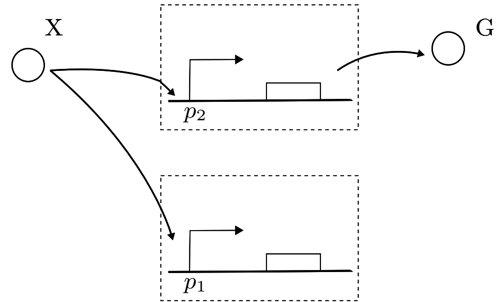

Figure 3: Transcription factor X regulates the production of protein G, while also binding to non-regulatory binding site p_1.

and $k_{\text{on}2}, k_{\text{off}2}$ are the binding/unbinding rate constants between the transcription factor X and the promoters p_1 and p_2, β is the production rate of the protein G, and δ_1, δ_2 are the decay rate constants of X and G respectively, which includes both degradation and dilution. The total amount of each promoter is conserved, and hence we can write $P_{t1} = p_1 + C_1$ and $P_{t2} = p_2 + C_2$. Denote by Ω the cell volume, and let $\Omega = 1$ for simplicity. Then, the chemical Langevin equations for the system can be written as

$$\begin{aligned}
\frac{dX}{dt} &= k(t) - \delta_1 X - k_{\text{on}1} X(P_{t1} - C_1) + k_{\text{off}1} C_1 - k_{\text{on}2} X(P_{t2} - C_2) + k_{\text{off}2} C_2 \\
&\quad + \sqrt{k(t)}\Gamma_1 - \sqrt{\delta_1 X}\Gamma_2 - \sqrt{k_{\text{on}1} X(P_{t1} - C_1)}\Gamma_3 + \sqrt{k_{\text{off}1} C_1}\Gamma_4 \\
&\quad - \sqrt{k_{\text{on}2} X(P_{t2} - C_2)}\Gamma_5 + \sqrt{k_{\text{off}2} C_2}\Gamma_6, && (45)\\
\frac{dC_1}{dt} &= k_{\text{on}1} X(P_{t1} - C_1) - k_{\text{off}1} C_1 + \sqrt{k_{\text{on}1} X(P_{t1} - C_1)}\Gamma_3 - \sqrt{k_{\text{off}1} C_1}\Gamma_4, && (46)\\
\frac{dC_2}{dt} &= k_{\text{on}2} X(P_{t2} - C_2) - k_{\text{off}2} C_2 + \sqrt{k_{\text{on}2} X(P_{t2} - C_2)}\Gamma_5 - \sqrt{k_{\text{off}2} C_2}\Gamma_6, && (47)\\
\frac{dG}{dt} &= \beta C_2 - \delta_2 G + \sqrt{\beta C_2}\Gamma_7 - \sqrt{\delta_2 G}\Gamma_8, && (48)
\end{aligned}$$

where Γ_i are independent white noise processes. We assume that the binding between the transcription factor X and the promoters are weak, giving $P_{t1} \gg C_1$ and $P_{t2} \gg C_2$. Therefore, we can write the system (45) - (48) in the form

$$\begin{aligned}
\frac{dX}{dt} &= k(t) - \delta_1 X - k_{\text{on}1} X P_{t1} + k_{\text{off}1} C_1 - k_{\text{on}2} X P_{t2} + k_{\text{off}2} C_2 + \\
&\quad \sqrt{k(t)}\Gamma_1 - \sqrt{\delta_1 X}\Gamma_2 - \sqrt{k_{\text{on}1} X P_{t1}}\Gamma_3 + \sqrt{k_{\text{off}1} C_1}\Gamma_4 - \sqrt{k_{\text{on}2} X P_{t2}}\Gamma_5 + \sqrt{k_{\text{off}2} C_2}\Gamma_6, \\
\frac{dC_1}{dt} &= k_{\text{on}1} X P_{t1} - k_{\text{off}1} C_1 + \sqrt{k_{\text{on}1} X P_{t1}}\Gamma_3 - \sqrt{k_{\text{off}1} C_1}\Gamma_4, \\
\frac{dC_2}{dt} &= k_{\text{on}2} X P_{t2} - k_{\text{off}2} C_2 + \sqrt{k_{\text{on}2} X P_{t2}}\Gamma_5 - \sqrt{k_{\text{off}2} C_2}\Gamma_6, \\
\frac{dG}{dt} &= \beta C_2 - \delta_2 G + \sqrt{\beta C_2}\Gamma_7 - \sqrt{\delta_2 G}\Gamma_8.
\end{aligned}$$

We have that the binding/unbinding reactions are much faster than protein production and decay [22], and thus we can write $\epsilon = \delta_1/k_{\text{off}1}$, where $\epsilon \ll 1$. Letting $k_{d1} = k_{\text{off}1}/k_{\text{on}1}$, $k_{d2} = k_{\text{off}2}/k_{\text{on}2}$, and $a = k_{\text{off}2}/k_{\text{off}1}$ we have $k_{\text{on}1} = \delta_1/(\epsilon k_{d1})$, $k_{\text{on}2} = a\delta_1/(\epsilon k_{d2})$, $k_{\text{off}1} = \delta_1/\epsilon$, and $k_{\text{off}2} = a\delta_1/\epsilon$. Then, with the change of variable $y = X + C_1 + C_2$, we can take the system into standard singular perturbation form

$$\begin{aligned}
\frac{dy}{dt} &= k(t) - \delta_1(y - C_1 - C_2) + \sqrt{k(t)}\Gamma_1 - \sqrt{\delta_1(y - C_1 - C_2)}\Gamma_2,\\
\epsilon\frac{dC_1}{dt} &= \frac{\delta_1}{k_{d1}}(y - C_1 - C_2)P_{t1} - \delta_1 C_1 + \sqrt{\epsilon}\sqrt{\frac{\delta_1}{k_{d1}}(y - C_1 - C_2)P_{t1}}\Gamma_3 - \sqrt{\epsilon}\sqrt{\delta_1 C_1}\Gamma_4,\\
\epsilon\frac{dC_2}{dt} &= \frac{a\delta_1}{k_{d2}}(y - C_1 - C_2)P_{t2} - a\delta_1 C_2 + \sqrt{\epsilon}\sqrt{\frac{a\delta_1}{k_{d2}}(y - C_1 - C_2)P_{t2}}\Gamma_5 - \sqrt{\epsilon}\sqrt{a\delta_1 C_2}\Gamma_6,\\
\frac{dG}{dt} &= \beta C_2 - \delta_2 G + \sqrt{\beta C_2}\Gamma_7 - \sqrt{\delta_2 G}\Gamma_8.
\end{aligned}$$

This system does not satisfy the sufficient conditions for existence of a unique solution in [16], and we note that the existence of a unique, well-defined solution for chemical Langevin equations is an ongoing research question [23, 24]. Therefore, in this case, we choose parameter conditions that give sufficiently large molecular counts, in order to increase the probability that the argument of the square-root term remains positive. In this example, the argument of the square-root remained positive for all the simulation runs performed and used to numerically determine the sample means.

Setting $\epsilon = 0$, we obtain the function $\gamma_1(y,t) = [\gamma_1(y,t)_1, \gamma_1(y,t)_2]^T$ in the form

$$\begin{aligned}
C_1 &= \frac{\frac{P_{t1}}{k_{d1}}y}{\frac{P_{t1}}{k_{d1}} + \frac{P_{t2}}{k_{d2}} + 1} = \gamma_1(y,t)_1,\\
C_2 &= \frac{\frac{P_{t2}}{k_{d2}}y}{\frac{P_{t1}}{k_{d1}} + \frac{P_{t2}}{k_{d2}} + 1} = \gamma_1(y,t)_2.
\end{aligned}$$

Then, to obtain the function $g(y,t)$ we consider the equation

$$g(y,t)g(y,t)^T B_2^T + B_2 g(y,t)g(y,t)^T = -\Lambda(y, \gamma_1(y,t), t, 0), \tag{49}$$

where B_2 and Λ are given by

$$B_2 = \begin{bmatrix} -\delta_1\frac{P_{t1}}{k_{d1}} - \delta_1 & -\delta_1\frac{P_{t1}}{k_{d1}} \\ -a\delta_1\frac{P_{t2}}{k_{d2}} & -a\delta_1\frac{P_{t2}}{k_{d2}} - a\delta_1 \end{bmatrix},$$

$$\Lambda = \begin{bmatrix} \frac{2\delta_1\frac{P_{t1}}{k_{d1}}y}{\frac{P_{t1}}{k_{d1}} + \frac{P_{t2}}{k_{d2}} + 1} & 0 \\ 0 & \frac{\frac{2a\delta_1 P_{t2}}{k_{d2}}y}{\frac{P_{t1}}{k_{d1}} + \frac{P_{t2}}{k_{d2}} + 1} \end{bmatrix}.$$

The eigenvalues of B_2 are given by $-\delta_1$ and $-\frac{\delta_1(k_{d2}P_{t1} + k_{d1}(k_{d2} + P_{t2})}{k_{d1}k_{d2}}$ where the parameters $\delta_1, k_{d1}, P_{t2}, k_{d2}, P_{t2}$ are positive. Therefore, we have that the matrix B_2 is Hurwitz. Then, solving the set of linear equations in (49), we find that the matrix $g(y,t)g(y,t)^T$ is given by

$$\frac{1}{(1+\frac{P_{t1}}{k_{d1}}+\frac{P_{t2}}{k_{d2}})^2}\begin{bmatrix} \frac{P_{t1}}{k_{d1}}(1+\frac{P_{t2}}{k_{d2}})y & -\frac{P_{t1}}{k_{d1}}\frac{P_{t2}}{k_{d2}}y \\ -\frac{P_{t1}}{k_{d1}}\frac{P_{t2}}{k_{d2}}y & \frac{P_{t2}}{k_{d2}}(1+\frac{P_{t1}}{k_{d1}})y \end{bmatrix}$$

and therefore we have

$$\begin{aligned} & g(y,t) \\ &= \frac{L}{\sqrt{S+2\sqrt{\Delta}}}\begin{bmatrix} \frac{P_{t1}}{k_{d1}}(1+\frac{P_{t2}}{k_{d2}})y+\sqrt{\Delta} & -\frac{P_{t1}}{k_{d1}}\frac{P_{t2}}{k_{d2}}y \\ -\frac{P_{t1}}{k_{d1}}\frac{P_{t2}}{k_{d2}}y & \frac{P_{t2}}{k_{d2}}(1+\frac{P_{t1}}{k_{d1}})y+\sqrt{\Delta} \end{bmatrix}, \end{aligned}$$

where $S = (\frac{P_{t1}}{k_{d1}}+\frac{P_{t2}}{k_{d2}}+2\frac{P_{t1}}{k_{d1}}\frac{P_{t2}}{k_{d2}})y$, $\Delta = \frac{P_{t1}}{k_{d1}}\frac{P_{t2}}{k_{d2}}(1+\frac{P_{t1}}{k_{d1}}+\frac{P_{t2}}{k_{d2}})y^2$ and $L = \frac{1}{(1+\frac{P_{t1}}{k_{d1}}+\frac{P_{t2}}{k_{d2}})}$.

Then, the reduced system is given by

$$\begin{aligned} \frac{dy}{dt} &= k(t) - \delta_1 Ly + \sqrt{k(t)}\Gamma_1 - \sqrt{\delta_1 Ly}\Gamma_2, \\ \frac{dG}{dt} &= \beta\frac{P_{t2}}{k_{d2}}Ly - \delta_2 G + \sqrt{\beta\frac{P_{t2}}{k_{d2}}Ly}\Gamma_7 - \sqrt{\delta_2 G}\Gamma_8, \\ C_1 &= \frac{P_{t1}}{k_{d1}}Ly + \frac{L\left(\left(\frac{P_{t1}}{k_{d1}}(1+\frac{P_{t2}}{k_{d2}})y+\sqrt{\Delta}\right)N_1 - \frac{P_{t1}}{k_{d1}}\frac{P_{t2}}{k_{d2}}yN_2\right)}{\sqrt{S+2\sqrt{\Delta}}}, \\ C_2 &= \frac{P_{t2}}{k_{d2}}Ly + \frac{L\left(-\frac{P_{t1}P_{t2}}{k_{d1}k_{d2}}yN_1 + \left(\frac{P_{t2}}{k_{d2}}(1+\frac{P_{t1}}{k_{d1}})y+\sqrt{\Delta}\right)N_2\right)}{\sqrt{S+2\sqrt{\Delta}}}, \end{aligned}$$

where N_1 and N_2 are standard normal random variables.

Fig. 4 shows the error in the moments between the fast variable dynamics of the original system and that of the approximation. The simulations are performed using the Euler-Maruyama method for stochastic differential equations and the moments are calculated using 500,000 simulation runs.

Remark: The above example is performed for illustration purposes and shows how the reduction approach can be applied to obtain the fast variable approximation. We also note that calculating the function $g(\bar{x},t)$ from $g(\bar{x},t)g(\bar{x},t)^T$ obtained through (16) may be challenging for systems with high dimension. However, in many applications, the analysis typically requires the calculation of statistical properties such as the mean and the variance, which can be directly calculated from the reduction approach using the functions $\gamma(x,t)$ and $g(\bar{x},t)g(\bar{x},t)^T$ in equations (7) and (16), which can be readily obtained.

6 Conclusion

We considered a class of singularly perturbed stochastic differential equations where the drift terms are linear and diffusion coefficients are nonlinear functions of the state variables.

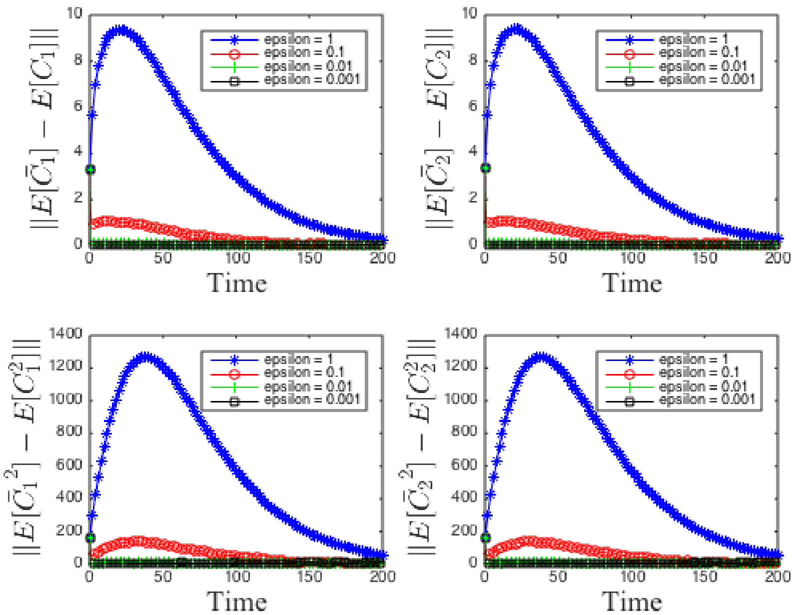

Figure 4: Errors in the first and second order moments. The parameters used are $k(t) = 10,\ \delta_1 = 0.1,\ \delta_1 = 1,\ k_{d1} = 1000,\ k_{d2} = 1000,\ P_{t1} = 1000,\ P_{t2} = 1000,\ \beta = 1,\ y(0) = 70,\ G(0) = 60,\ C_1(0) = 20$ and $C_2(0) = 20$.

Building on the results in our previous paper [10], where we obtained a reduced system that approximates the slow variable dynamics, in this work, we obtained an approximation for the fast variable, when the time-scale separation is large. This result allows the derivation of a reduced-order system with approximations for both slow and fast dynamics, which is useful in many applications. In particular, biomolecular systems consist of variables affected by both slow and fast reactions, which can be represented in singular perturbation form after a coordinate transformation. This approach could be used to analyze the statistical properties of such systems where the variables of interest are affected by both slow and fast dynamics.

In future work, we aim at extending this analysis to systems with nonlinear drift terms.

Acknowledgement

This work was funded by AFOSR grant # FA9550-14-1-0060.

References

[1] N. Herath and D. Del Vecchio. Model reduction for a class of singularly perturbed stochastic differential equations : Fast variable approximation. In *American Control Conference (ACC), 2016*.

[2] H. K. Khalil. *Nonlinear systems*, volume 3. Prentice hall Upper Saddle River, 2002.

[3] P. Kokotović, H. K. Khalil, and J. O'reilly. *Singular perturbation methods in control: analysis and design*, volume 25. Siam, 1999.

[4] G. Pavliotis and A. Stuart. *Multiscale methods: averaging and homogenization*, volume 53. Springer Science & Business Media, 2008.

[5] Y. Kabanov and S. Pergamenshchikov. *Two-scale stochastic systems: asymptotic analysis and control*, volume 49. Springer, 2003.

[6] N. Berglund and B. Gentz. *Noise-induced phenomena in slow-fast dynamical systems: a sample-paths approach*. Springer, 2006.

[7] C. Tang and T. Başar. Stochastic stability of singularly perturbed nonlinear systems. In *Decision and Control, 2001. Proceedings of the 40th IEEE Conference on*, volume 1, pages 399–404. IEEE, 2001.

[8] D. T. Gillespie. The chemical langevin equation. *The Journal of Chemical Physics*, 113(1):297–306, 2000.

[9] R. Z. Khasminskii. On the principle of averaging the itô stochastic differential equations. *Kybernetika (Prague)*, 4:260–279, 1968.

[10] N. Herath, A. Hamadeh, and D. Del Vecchio. Model reduction for a class of singularly perturbed stochastic differential equations. In *American Control Conference (ACC), 2015*, pages 4404 – 4410.

[11] N. Herath and D. Del Vecchio. Moment convergence in a class of singularly perturbed stochastic differential equations. In *Control Conference (AUCC), 2015 5th Australian*, pages 43–48.

[12] M. Contou-Carrere, V. Sotiropoulos, Y. N. Kaznessis, and P. Daoutidis. Model reduction of multi-scale chemical langevin equations. *Systems & Control Letters*, 60(1):75–86, 2011.

[13] D. Del Vecchio, A. J. Ninfa, and E. D. Sontag. Modular cell biology: retroactivity and insulation. *Molecular systems biology*, 4(1), 2008.

[14] A. Gyorgy and D. Del Vecchio. Retroactivity to the input in complex gene transcription networks. In *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, pages 3595–3601, Dec 2012.

[15] B. Øksendal. *Stochastic differential equations*. Springer, 2003.

[16] D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical finance*, 6(4):379–406, 1996.

[17] H. Kwakernaak and R. Sivan. *Linear optimal control systems*, volume 1. Wiley-interscience New York, 1972.

[18] S. Jayanthi, K. S. Nilgiriwala, and D. Del Vecchio. Retroactivity controls the temporal dynamics of gene transcription. *ACS synthetic biology*, 2(8):431–441, 2013.

[19] T. Lee and N. Maheshri. A regulatory role for repeated decoy transcription factor binding sites in target gene expression. *Molecular systems biology*, 8(1):576, 2012.

[20] D. Mishra, P. M. Rivera, A. Lin, D. Del Vecchio, and R. Weiss. A load driver device for engineering modularity in biological networks. *Nature biotechnology*, 32(12):1268–1275, 2014.

[21] M. Soltani, P. Bokes, Z. Fox, and A. Singh. Nonspecific transcription factor binding can reduce noise in the expression of downstream proteins. *Physical biology*, 12(5):055002, 2015.

[22] U. Alon. *An introduction to systems biology: design principles of biological circuits*. CRC press, 2006.

[23] J. Wilkie and Y. Wong. Positivity preserving chemical langevin equations. *Chemical Physics*, 353(1):132–138, 2008.

[24] D. Schnoerr, G. Sanguinetti, and R. Grima. The complex chemical langevin equation. *The Journal of chemical physics*, 141(2):024103, 2014.